ED 409 332 TM 026 775

AUTHOR	Nandakumar, Ratna; Roussos, Louis
TITLE	Validation of CATSIB To Investigate DIF of CAT Data.
PUB DATE	28 Mar 97
NOTE	28p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Adaptive Testing; *Computer Assisted Testing; *Item Bias; *Regression (Statistics); Sample Size; *Statistical Distributions; *Test Format
IDENTIFIERS	*CATSIB Computer Program; Item Bias Detection; Power (Statistics); *Type I Errors

ABSTRACT

This paper investigates the performance of CATSIB (a modified version of the SIBTEST computer program) to assess differential item functioning (DIF) in the context of computerized adaptive testing (CAT). One of the distinguishing features of CATSIB is its theoretically built-in regression correction to control for the Type I error rates when the distributions of the reference and focal groups differ on the intended ability. This phenomenon is also called impact. The Type I error rate of CATSIB with the regression correction (WRC) was compared with that of CATSIB without the regression correction (WORC) to see if the regression correction was indeed effective. Also of interest was the power level of CATSIB after the regression correction. The subtest size was set at 25 items, and sample size, the impact level, and the amount of DIF were varied. Results show that the regression correction was very useful in controlling for the Type I error, CATSIB WORC had inflated observed Type I errors, especially when impact levels were high. The CATSIB WRC had observed Type I error rates very close to the nominal level of 0.05. The power rates of CATSIB WRC were impressive. As expected, the power increased as the sample size increased and as the amount of DIF increased. Even for small samples with high impact rates, power rates were 64% or higher for high DIF levels. For large samples, power rates were over 90% for high DIF levels. (Contains 12 tables and 7 references.) (Author/SLD)

Validation of CATSIB to investigate DIF of CAT Data

Ratna Nandakumar
University of Delaware
and
Louis Roussos
Law School Admission Council

Note: Authors names are listed in alphabetical order.

Paper presented at the annual meeting of the American Educational Research Association, Chicago, March 28, 1997.

Abstract

This paper investigated the performance of CATSIB to assess DIF in the context of computerized adaptive testing. One of the distinguishing features of CATSIB is its theoretically built in regression correction to control for the Type I error rates when the distributions of the reference and focal groups differ on the intended ability. This phenomenon is also called *impact*, (d_T). It was of special interest to investigate the Type I error rate of CATSIB with the regression correction (WRC) with that of without the regression correction (WORC) to see if the regression correction was indeed effective. Also of interest was the power level of CATSIB after the regression correction. The matching subtest size was fixed to 25 items. Three factor were varied: the sample size, the impact level, and the amount of DIF (β). Three levels of examinee sample sizes were used : 250 examinees in the reference group and 250 in the focal group; 500 and 250; and 500 and 500. Three levels of d_T were used: 0, .5, and 1.0. Three β levels were used: no DIF, moderate DIF and high DIF. Examinee sample sizes and impact levels were completely crossed for the Type I error study using no DIF items; same was done with complete crossing of two levels of DIF (moderate and high) for the power study.

The results showed that the regression correction was very useful in controlling for the Type I error. CATSIB WORC had inflated observed Type I errors, especially when the impact levels were high ($d_T = 1$). Moreover, the number of rejection rates that were out of bounds increased as the sample size increased. The CATSB WRC, on the other hand, had observed Type I error rates very close to the nominal level of .05. The power rates of CATSIB WRC were impressive. As expected, the power increased as the sample size increased, and as the amount of DIF increased. Even for small samples with high impact levels the power rates were 64% or higher for high DIF levels. For large samples the power rates were over 90% for high DIF levels.

Currently, computerized testing and computerized adaptive testing (CAT) are increasingly becoming popular for standardized tests. For example, examinees can now opt for computerized adaptive tests of Scholastic Assessment Tests (SAT), Graduate Record Examination (GRE), and the Armed Services Vocational Aptitude Battery (ASVAB). In any testing situation, it is essential that items are fair to all subroups of examinees. In other words, if on an item, one subgroup is doing better than the other, although both subgroups are of equal ability, then such an item will be cause for a concern. This phenomenon is known as differential item functioning (DIF). A number of studies have been conducted using different methodologies to investigate DIF of paper-and-pencil tests. In a typical paper-and-pencil test, most DIF analyses involve matching examinees on the number-correct score, a traditional matching criterion. However, in the context of CAT, not all examinees take the same items, and sometimes not all take the same number of items. Hence one can not apply methods developed for paper-and-pencil tests to assess DIF of CAT items. Zwick, Thayer, and Wingersky (1994, 1995) have studied DIF of computerized adaptive test items using a modified version of the Mantel Haenszel methodology. Roussos (1996) has developed a modified version of SIBTEST, called CATSIB, to identify DIF items of an adaptive test using a matching criterion based on CAT ability estimates. Roussos in his simulation study, however, has investigated the Type I error performance of CATSIB using ability estimates of a paper-and-pencil test and not an adaptive test. The purpose of the present study was to investigate the performance of CATSIB to identify DIF items of adaptive tests through simulations . Both the Type I error rate and power level were of interest in the present study. CATSIB will be briefly described below followed by details of the simulation study, results and discussion.

CATSIB

In this methodology the conceptualization of DIF is based on the Shealy and Stout (1993) multidimensional model for DIF. That is, DIF is manifested through multidimensionality. Since the test is assumed to be measuring a unidimensional construct (θ), DIF is manifested through items tapping other abilities (ηs) in addition to the intended ability, θ. The subgroups for whom the DIF is being investigated are referred to as the reference (R) group and the focal (F) group. If the conditional distributions of η differ between R and F groups who are matched on the intended ability, then the item is said to exhibit DIF. The item being studied is referred to as the *studied item*. In paper-and-pencil tests the matching criterion used for comparing R and F groups is usually the total test score or the total score on the rest of the items of the test. In a CAT setting, the matching criterion is based on the ability estimates ($\hat{\theta}$) obtained from adaptively administered items. The null hypothesis for DIF is stated as

$$H_0 : \beta_{uni} = 0,$$

where β_{uni} is the uniform DIF of a specific item or a collection of items under investigation and is given by

$$\beta_{uni} = \int DIF(\theta) f(\theta) d\theta, \tag{1}$$

where

$$DIF(\theta) = P_R(\theta) - P_F(\theta)$$

is the difference in the probabilities of getting the studied item correct between R and F groups of examinees of ability θ, and $f(\theta)$ is an appropriate density function such as that for the combined R and F groups.

It is generally known that R and F groups differ on the intended ability (θ) distributions, known as *impact*, denoted by $d_T = \mu_{\theta_R} - \mu_{\theta_F}$, where μ_{θ_R} and μ_{θ_F} are the means of the ability distributions of R and F groups respectively. In such cases the use of $\hat{\theta}$s in the computation of DIF could lead to a high Type I error because $E_R[\theta|\hat{\theta}]$ can be much different from $E_F[\theta|\hat{\theta}]$. Therefore, in order to avoid the high Type I error rate CATSIB matches examinees on $\hat{E}[\theta|\hat{\theta}]$ denoted by $\hat{\theta}^*$ and is given by

$$\hat{E}[\theta|\hat{\theta}] = \hat{\theta}^* = \hat{\theta}_G + \hat{\rho}_G(\hat{\theta} - \bar{\hat{\theta}}_G), \text{ where G= R or F.}$$

$\bar{\hat{\theta}}_G$ is the mean of the distribution of ability estimates of group G. The reliability estimate $\hat{\rho}_G$ is given by

$$\hat{\rho}_G = 1 - \hat{\sigma}_e^2 / \hat{\sigma}_{\hat{\theta}}^2,$$

where $\hat{\sigma}_{\hat{\theta}}$ is the observed standard deviation of ability estimates of group G, and

$$\hat{\sigma}_e^2 = 1/\hat{E}_{\hat{\theta}}[\sum_i I_i(\hat{\theta})]$$

is the estimated standard deviation of the error distribution of group G. $I_i(\hat{\theta})$ is the Fisher information for item i at ability estimate $\hat{\theta}$ (Hambleton, Swaminathan, & Rogers, 1991, p. 91). In the present study true item parameters were used in computing $I_i(\hat{\theta})$.

The test statistic for testing the null hypothesis of no DIF is then given by

$$B = \frac{\hat{\beta}_{uni}}{\hat{\sigma}(\hat{\beta}_{uni})} \tag{2}$$

where

$$\hat{\beta}_{uni} = \hat{\beta} = \sum_{\hat{\theta}^*=\hat{\theta}^*_{min}}^{\hat{\theta}^*_{max}} [\hat{P}_R(\hat{\theta}^*) - \hat{P}_F(\hat{\theta}^*)]p(\hat{\theta}^*),$$

and $\hat{\sigma}(\hat{\beta}_{uni})$ is its estimated standard error, and $p(\hat{\theta}^*)$ is the proportion of R and F examinees at $\hat{\theta}^*$. The null hypothesis of no DIF is rejected at level α if the statistic B exceeds the $100(1-\alpha)$ percentile obtained from the standard normal table.

The estimate $\hat{\beta}$ serves as an index of the amount of DIF present in the item. For example, it is possible that an item may exhibit statistically significant DIF but the degree of DIF may not be practically meaningful in terms of how it affects the performance of examinees in the two groups. $\hat{\beta}$ can be very useful in assessing the degree of DIF practically. It can be seen from Equation 1 that $\hat{\beta}$ denotes the average difference in probabilities of correct response on the given item between R and F groups. When $\hat{\beta} = .05$, for example, the difference corresponds to a 5% difference, which can be considered as moderate DIF. When $\hat{\beta} = .1$, on the other hand, the difference in probabilities of correct response between groups corresponds to 10%, which may be considered as high DIF. This categorization could also be subjective depending upon the nature of item parameters. For example, a $\hat{\beta} = .05$ may be a big difference if the item is a difficult one as opposed to an easy item. Throughout this paper a $\hat{\beta} = .05$ is considered a moderate DIF item and a $\hat{\beta} = .1$ is considered a high DIF item.

Unique features of CATSIB are that it has a theoretically based correction for adjusting the $\hat{\theta}$s of the R and F groups due to differences on the intended ability distributions (impact); and the advantage of assessing DIF either for a single item or a collection of items.

The Simulation Study

In large testing companies DIF analyses are routinely carried out during the pretesting process. The simulation study will therefore mimic a pretest scenario. The objectives of the proposed study are to assess the Type I error rate and power level of CATSIB for computerized adaptive tests.

The matching subtest length was fixed to 25 items, which is typical of a standard CAT. That is, each examinee was administered 25 items adaptively from a pool of 1000 items. Three factors were varied in this study: the sample size, the impact level (d_T), and the amount of DIF (β). Three different combinations of examinee sample sizes were selected: n_R=250, n_F=250; n_R=500, n_F=250; and n_R=500, n_F=500; where n_R and n_F denote the sample sizes in the R and F groups respectively. Three different impact levels were used: 0, .5, and 1. These three d_T levels correspond to differences in the means of the ability distributions of R and F groups that are 0, .5 standard deviation, and 1 standard deviation apart. The sample sizes and impact levels were completely crossed resulting in 9 combinations of sample size and d_T level combinations. Three DIF levels were used: no DIF ($\beta = 0$), moderate DIF ($\beta = .05$), and high DIF ($\beta = .10$).

The means of ability distributions for R and F groups were determined in such a way that their weighted average was centered around mean difficulty (which is equal to 0) with the desired impact level. This was accomplished by solving for $\mu_{\theta R}$ and $\mu_{\theta F}$ from the following two equations:

$$\alpha_R \mu_{\theta R} + \alpha_F \mu_{\theta F} = 0$$

$$\mu_{\theta R} - \mu_{\theta F} = d_T$$

where $\alpha_R = \frac{n_R}{n_r + n_F}$; $\alpha_F = \frac{n_F}{n_r + n_F}$. The standard deviations of ability distributions were each set to 1.

In generating the item parameters of the item pool for the matching subtest, the goal was to generate parameters that closely resembled those of real parameters. Upon observing the descriptive properties of item pools of a major testing company it was found that the distribution of item discriminating parameters ranged from .5 to 1.7 and followed a positively skewed distribution; while item difficulty parameters ranged from -2 to 2 and followed the standard normal distribution. The discrimination parameters were therefore generated from a lognormal distribution, and difficulty parameters were generated to follow the standard normal distribution. The lower asymptote was independently generated from a uniform distribution to range between .12 and .22 to approximate the real parameters. The distribution of item parameters are described below.

$$\text{a} \sim \text{lognormal } (-.357, .25) \text{ for b} \leq -1 \text{ with range } .4 \leq \text{a} \leq 1.1$$

$$\text{a} \sim \text{lognormal } (-.223, .34) \text{ for b} > -1 \text{ with range } .4 \leq \text{a} \leq 1.7$$

$$\text{b} \sim N(0,1) \text{ with range } -3 \leq \text{b} \leq 3$$

$$\text{c} \sim \text{U}(.17) \text{ with range } .12 \leq \text{c} \leq .22$$

The item parameters of DIF items were systematically varied according to the discrimination and difficulty parameters of items. DIF was introduced through difference in the difficulty parameters of the studied item between R and F groups using the following model for DIF:

$$\text{DIF} = \hat{\beta} = \int [P_R(\theta) - P_F(\theta)] f(\theta) d\theta, \tag{3}$$

where

$$P_{iG}(\theta) = c_i + \frac{1 - c_i}{1 + exp(-1.7(a_i(\theta - b_{iG})))}, G = R or F \tag{4}$$

For each level of DIF (β: 0, .05 or .1), five to six items were selected such that they are a combination of items of varying difficulty and discrimination parameters. For example, when DIF=.05, five items were chosen as follows: low discrimination($a = .4$) and low difficulty ($b \approx -1.5$); low discrimination($a = .4$) and high difficulty ($b \approx 1.5$); medium discrimination($a = .8$) and medium difficulty ($b \approx 0$); high discrimination($a = 1.0$) and low difficulty ($b \approx -1.5$); high discrimination($a = 1.4$) and high difficulty ($b \approx 1.5$). The difficulty parameters of DIF items were adjusted between R and F groups so that items exhibit the desired level of DIF. That is, b_i for R and F groups were determined by trial and error approach so that they satisfy Equation 3 for a given level of DIF with the constraint that b_R and b_F average to -1.5, 0, or 1.5. There were in total 16 DIF items: 6 at DIF=0 level, 5 each at DIF levels of .05 and .1. The DIF items along with their parameters are listed in the first four columns of Tables 1 to 9. The first six items with DIF=0 are Type I error items. That is, for these items Type I error performance of CATSIB was investigated. For the remaining 10 items power performance of CATSIB was investigated. Items 6 to 11 were moderate DIF items ($\beta = .05$), and items 12 to 16 were high DIF items ($\beta = .10$).

The CAT Procedure

For a given combination of examinees' sample size (n_R and n_F) and impact level (d_T), examinees of R and F groups were simulated from their respective distributions[1] Each examinee in each of the groups R and F was administered a fixed length test of 25-items adaptively from the item pool of 1000 items. The ability estimates of examinees were determined using the standard maximum-information CAT design described as follows.

The ability scale from -2.25 to 2.25 was divided into 37 equal intervals in increments of 0.125. For each item i, item information, $I_i(\theta)$, was computed at each of the θ values

[1]The R group examinees were simulated from N($\mu_{\theta R}$,1) distribution and the F group examinees were simulated from N($\mu_{\theta F}$,1) distribution.

using the following formula (Hambleton, Swaminathan, & Rogers, 1991, p. 91):

$$I_i(\theta) = \frac{2.89a_i^2(1-c_i)}{[c_i + exp(1.7a_i(\theta - b_i))][1 + exp(-1.7a_i(\theta - b_i))]^2}$$

where a_i, b_i, and c_i denote discrimination, difficulty, and lower asymptote parameters of item i respectively. At each θ level the matching subtest pool of items were sorted according to the item information values from lowest to highest and saved in a separate table. This table was used during the simulations to select items with the highest information at a given θ-level.

To prevent items from becoming overexposed, an exposure control method was incorporated (Kingsbury & Zara, 1989). Accordingly, the first item to be administered to a simulated examinee was randomly selected from the 10 items with highest information values at $\theta = 0$ (the starting value for all simulees). The second item was randomly selected from the 9 best items at the new estimate of θ. The third item was randomly selected from the 8 best items, and so on until, beginning with the 10th item, the item with the highest information was selected (unless, of course, the item had already been administered to that simulee, in which case, the next best item was selected).

After administering each item in this manner, to each examinee, the simulated examinee's response (right/wrong) was determined, and the simulated examinee's estimated ability, $\hat{\theta}$, was updated using Owen's Bayesian sequential scoring (Owen, 1969). After all 25 items were administered, a Bayesian modal score was calculated and was used as the final ability estimate ($\hat{\theta}$).

After obtaining $\hat{\theta}$s for all examinees in R and F groups, they were formed into subgroups based on their $\hat{\theta}$s. That is, examinees with similar scores on $\hat{\theta}$ in R and F groups were placed into one subgroup. The minimum number of examinees in each subgroup (cell size) was taken to be 3 for R and F groups. The number of subgroups was adjusted so that over 90% of examinees were included in the analysis. There were 20 to 80 subgroups depending upon the sample size and d_T levels. For example, when n_R=n_F=250, and d_T=1.0 there were typically about 20 subgroups. Within each subgroup the regression correction was applied and examinees' corrected ability estimates $\hat{\theta}^*$s were obtained. The examinees were then reclassified into subgroups based on their corrected ability estimates.

The DIF items were then nonadaptively administered one at a time to all the examinees. After each administration of a DIF item, the bias estimate $\hat{\beta}$, and the statistic B were computed and tested for the presence of DIF using a right-tailed test, a left-tailed test, and a two-tailed test. The left tail test involves rejecting the null hypothesis of no DIF at level $\alpha = .05$ if the computed z-statistic is less than -1.645. The right tail test rejects if the computed z-statistic is greater than 1.645; and the two-tailed test rejects if $|z| \geq 1.96$.

For a given combination of sample size and the d_T level, this process, starting from the simulation of $\hat{\theta}$s, was replicated 400 times for each DIF item (16 DIF items in total). The average DIF estimate $\bar{\hat{\beta}}$ over 400 replications, its standard error, the rejection rates for the right-tailed test (RT rr), the left-tailed test (LT rr), and the two-tailed test (2T rr) were computed for CATSIB with regression correction and for CATSIB without regression correction. These results are reported in Tables 1-9.

Results

The Type I Error Study

The results of the simulation study are reported in Tables 1 through 9. The top half of each table shows the results of CATSIB with the regression correction (WRC), and the bottom half shows results for CATSIB without the regression correction (WORC). The first column, in each table, shows the item number, columns two to four show item parameters: the discrimination parameter, the difficulty parameter for R group and the difficulty parameter for F group respectively. The fifth column shows the estimated DIF averaged over 400 trials ($\bar{\hat{\beta}}$), and the sixth column shows its standard error. Columns seven to nine show the rejection rates for: the left tail test (LT rr), the right tail test (RT rr), and the two-tailed test (2T rr) respectively. The last two columns give rejection rates based on two criteria. Namely, rejection of H_0 with low DIF (the computed $|z| \geq 1.96$ and $\hat{\beta} > .05$) in column 9; and rejection of H_0 with high DIF (the computed $|z| \geq 1.96$ and $\hat{\beta} > .1$) in column 10. Among the items (rows), the first six items are no DIF items ($\beta = .0$). Therefore only the Type-I error rate was of interest for these items. The remaining 10 items are DIF items and hence only the power rate was of interest for these items. Summaries of Type I error study are reported in Tables 10, 11 and in Figure 1; summaries of power study are reported in Table 12.

For no DIF items (1 to 6), since α=.05, the 95% confidence intervals for α across all tables are .0471 and .0529 [2]. The observed average Type I error rates for items 1 to

[2]Across all tables, there are 21,600 trials (400x6x9) for each of the tails. Therefore the standard error becomes very small (.00148).

6 across all tables together were .0487, .0518, and .0509 for the LT rr, RT rr, and 2T rr respectively for the CATSIB WRC. They all fell within the 95% confidence bounds, suggesting they all are not significantly different from .05. The corresponding rates for CATSIB WORC were, .0394, .0683, and .0540, and all fell outside the confidence bound, suggesting they are all significantly different from .05. The summaries of Type I error rates by sample size and d_T level are reported in Table 10. For this table the observed values of α are expected to fall between .0413 and .0587 95% of the time. It can be seen from Table 10 that while the observed average rejections for the CATSIB WRC were generally within chance of the nominal level α, they were either inflated or deflated for the CATSIB WORC, with majority (75%) of them out of bounds in the presence of impact. For CATSIB WORC there was a consistent pattern of serious deflation of Type I error rates for the LT across all levels of sample sizes, and a corresponding serious inflation of Type I error rates for the RT.

For each entry in Tables 1 to 9, for items 1 to 6, the confidence bounds (based on 400 trials) for the observed α fall between .03 and .07. It can be seen from these tables that some of the rejection rates were out of bounds, below .03 or above .07. Table 11 lists the number of rejection rates that were out of bounds by sample size and d_T level for both CATSIB WRC and CATSIB WORC. Each cell entry is the number of rejection rates out of bounds out of 18 cases (6itemsx3tests). Based on exact binomial probabilities, out of 18 test, one expects 0, 1, or 2 rejections due to chance alone about 95% of the time. From Table 11 it can be seen that, for the CATSIB WORC, at each sample size, as the level of impact increases, the number of tests out of bounds also increases. Figure 1 shows a graphical display of these results. It is evident that as the impact level increases, the Type I error inflation also increases, but the degree of inflation is much steeper for CATSIB without regression correction. For the large impact level (d_T=1.0), the number of rejection rates out of bounds for CATSIB WORC were grossly inflated over the chance levels, while for the CATSIB WRC the inflation was only slightly more than the chance level.

The estimated DIF ($\bar{\hat{\beta}}$) was close to 0 for CATSIB WRC throughout as expected and its standard error was also small, with even smaller values as the sample size got larger. Under H_0, the distribution of the statistic B, of Equation 2, follows a normal distribution with mean 0 and standard deviation 1. The empirical distribution of B in the case of CATSIB WRC followed the theoretical distribution. In the case of CATSIB WORC, however, the mean of the distribution of B was inflated. The regression correction is therefore extremely useful in controlling for the Type I error inflation, and in maintaining theoretical properties of the distribution of B under the null hypothesis.

The Power Study

For this study, only results of CATSIB WRC will be analyzed because the results of CATSIB WORC are distorted by the statistical bias evident in its Type I error inflation (and as evident in power results of Tables 1 to 9). Therefore, only top parts of Tables 1 to 9 are referred. In Tables 1 to 9, items 7 to 11 are of moderate DIF (MDIF) and 12 to 16 are of high DIF (HDIF) items. Within each group, items are mixed in terms of low and high difficulty and discrimination parameters. Since the DIF was induced for F group examinees, only rejection rates for the right-tail (RT rr) and the two-tail (2T rr) were of interest. For a statistical test exhibiting good power, the rejection rates should increase as the amount of DIF increases and/or as the sample size increases. This is true with CATSIB (WRC) for items in all tables. Overall, as the DIF level increased from moderate ($\beta = .05$) to high ($\beta = .1$), the average power rates went up from 45% to 87% for the RT rr and from 34% to 80% for the 2T rr. Similarly, as the sample size increased from 500 (n_R=250, n_F=250) to 1000 (n_R=500, n_F=500), the power increased from 58% to 76% for the RT rr and from 48% to 68% for the 2T rr. Table 12 provides detailed summaries of power rates by d_T levels and sample sizes. The top half of Table 12 provides summaries for the RT rr and the bottom half of Table 12 provides summaries for the 2T rr. As can be seen from these results, at every level of d_T, the rejection rates generally increased by almost 100% (or higher in some cases) as the level of DIF increased from moderate to high. Similarly, as expected, as the sample size increased, the power also increased consistently. Naturally, the increase was much more for MDIF than for HDIF. Even for small samples such as n_R=250, n_F=250, the power was remarkably high as the DIF levels reached high.

In all cases, the amount of estimated DIF ($\bar{\hat{\beta}}$)was close to the true values (.05 for items 7 to 11, and .1 for item 12 to 16) and the standard errors were all small and showed only the very slightest increase as the level of impact went from .5 to 1. Item parameters have also played a role in the power rates. For example, items 10 and 15 with moderate discrimination and low difficulty had relatively higher power than other items in every Table (1 to 9).

The last two columns of Tables 1-9 show the rejection rates for moderate DIF and high DIF values. For example, for items 12 to 16 the DIF level was high ($\beta = .1$). Accordingly, the average estimated DIF ($\bar{\hat{\beta}}$), across 400 replications, was approximately equal to .1, which is also equal to the mean of the distribution of $\hat{\beta}$s. Since the mean of the distribution of $\hat{\beta}$s is equal to .1, one expects about 50% of the $\hat{\beta}$s to be greater than .1, and about 50% less than .1. It can be seen form the last column of Tables 1 to 9 that for items 12 to 16 the rejection rate was approximately 50%. That is, all $\hat{\beta}s \geq .1$

lead to rejections. Based on the reported standard errors for the smallest sample sizes (250, 250), a HDIF item having $\beta = .17$ or higher would be detected with 95% or more power no matter what its difficulty or discrimination levels are. Moreover, for a high dscriminating and easy item, such as item 10, HDIF item with β of only .146 or higher will be detected with 95% or more power. Similarly, for items 7 to 11, the average estimate of DIF ($\overline{\hat{\beta}}$), across 400 replications, was approximately equal to .05, which is also equal to the mean of the distribution of $\hat{\beta}$s for these items. Thus one expects about 50% of the $\hat{\beta}$s to be greater than .05, and about 50% less than .05. Unlike high DIF items, the rejection rates of items that have exhibited only moderate DIF are less than 50% (column 11 of Tables 1 to 9) for small sample sizes, and reached close to 50% as the sample sizes increased. Based on the reported standard errors for the smallest sample sizes (250, 250), a MDIF item having $\beta = .08$ or higher would be detected with 50% or more power no matter what its difficulty or discrimination levels are. Moreover, for a high dscriminating and easy item, such as item 10, MDIF item with β of only .055 or higher will be detected with 50% or more power.

In summary, the simulation results have shown that CATSIB with regression correction has performed desirably and the regression correction was very effective in controlling for the Type I error. The power results were noteworthy, in that, for sample sizes as small as 250 examinees in each group in combination with impact levels as large as one standard deviation, the Type I error rates have generally been within expected boundaries with good power rates for moderate DIF and very high power rates for high DIF.

Summary and Discussion

CATSIB is a DIF detecting procedure developed by Roussos (1996) for the computerized adaptive tests. Roussos investigated its performance for the paper-and-pencil tests and found the regression correction to be effective in controlling for the Type I error rate. Although the regression correction has been found to be very effective for paper-and-pencil tests, its effectiveness for computerized adaptive tests was not known. The present study was the first one designed to investigate the Type I error and power performance of CATSIB WRC for the computerized adaptive tests. Parameters of this study were designed to be as realistic as possible. The matching subtest size was fixed to 25 items. Three factor were varied: the sample size, the impact level, and the amount of DIF. Three levels of examinee sample sizes were used : 250 examinees in the reference group and 250 in the focal group; 500 and 250; and 500 and 500. Three levels of impact levels, d_T ,were used: 0, .5, and 1.0. Three β levels were used: no DIF, moderate DIF and high DIF. Examinee sample size and impact levels were completely crossed for the Type I error study using no DIF items; and same was done with complete crossing of two DIF levels (moderate and high) for the power study. 16 different DIF items were used in this study: the first 6 were no DIF items ($\beta = 0$), the next five were moderate DIF items ($\beta = .05$), and the last five were high DIF items ($\beta = .1$). Within each group, items were mixed with respect to their discrminating and difficulty parameters: low discrimination and low difficulty; low discrimination and high difficulty; medium discrimination and medium difficulty; high discrimination and low difficulty; high discrimination and high difficulty. For each combination of factors, and for each DIF item, 400 replications were simulated and tested for presence of DIF. The mean estimate of DIF, its standard error, and rejection rates for the LT, RT, and 2T were computed. In addition, the percentage of items that exhibited moderate DIF and led to rejection were computed (MDIF). Similarly, the percentage of items that exhibited high DIF and led to rejection were computed (BDIF).

The Type I error study compared the Type I error rates of CATSIB WRC with CATSIB WORC. The results showed that the regression correction was very effective in controlling for the Type I error. CATSIB WORC had inflated observed Type I errors, especially when the impact levels were high ($d_T = 1$). Also the rejection rates that were out of bounds increased as the sample size increased. The CATSB WRC, on the other hand, had observed Type I error rates very close to the nominal rate of .05. The power rates of CATSIB WRC were impressive. As expected, the power increased as the sample size increased, and as the amount of DIF increased. Even for small samples with high impact levels the power rates were 64% or higher when the degree of DIF was high. For large samples the power rates were over 90% for high DIF levels.

The results of this study have shown that CATSIB is a practical and a reliable statistical procedure for detecting DIF of computerized adaptive tests. It has performed satisfactorily under the conditions controlled in this study and therefore shows high potential for operational use. The Type I error rates and power rates were as good as one could hope for. For items with moderate DIF, although the power was lower for small samples with combination of high impact levels, it steadily increased as the sample sizes increased. That is, not surprisingly, one needs high sample sizes when comparing diverse groups with respect to their ability distributions, in order to detect moderate amounts of DIF. The use of $\hat{\beta}$ for assessing the amount of DIF can be very useful in applications. Since $\hat{\beta}$, the estimated degree of DIF, is the difference in probabilities of correct responses between R and F groups on the studied item, this index can be used in judgments about whether to keep this item in the pool or not for future administrations.

This study is limited in terms of factors varied. Further studies are needed to investigate its performance in more situations under varied conditions. For example, in this study, true (rather than estimated) item parameters were used in computing the information function. It will be of interest to study its performance with estimated parameters. Also CATSIB's performance on real data must be investigated. Another important practical issue is the cut-off values for $\hat{\beta}$ that are practically meaningful.

References

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education*, 2, 359-375.

Owen, R. J. (1969) Bayesian approach to tailored testing. (Research Bulletin 69-92). Princeton, NJ: Educational Testing Service.

Roussos, L. (1996, June). A type I error rate study of a modified SIBTEST DIF procedure with potential application to computerized-adaptive tests. Paper presented at the annual meeting of the Psychometric Society, Banff, Alberta, Canada.

Shealy, R., & Stout, W. (1993a). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*, 159-194.

Zwick, R., Thayer, D.T., & Wingersky, M. (1994). A simulation study of methods for assessing differential item functioning in computer-adaptive tests, Applied Psychological Measurement, 18, 121-140.

Zwick, R., Thayer, D.T., & Wingersky, M. (1995). Effect of Rasch calibration on ability and DIF estimation in computer-adaptive tests. Journal of Educational Measurement, 4, 341-363.

Table1
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=250, n-F=250, d-T=0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.40	-1.500	-1.500	0.001	0.0018	0.0275	0.0500	0.0425	0.0425	0.0075
2	0.40	1.500	1.500	-0.003	0.0023	0.0500	0.0425	0.0525	0.0525	0.0325
3	0.80	0.000	0.000	-0.001	0.0022	0.0600	0.0575	0.0475	0.0475	0.0225
4	1.00	-1.500	-1.500	-0.001	0.0014	0.0325	0.0475	0.0425	0.0400	0.0000
5	1.40	1.500	1.500	0.004	0.0019	0.0375	0.0575	0.0500	0.0500	0.0050
6	1.40	-1.500	-1.500	0.002	0.0012	0.0575	0.0550	0.0700	0.0450	0.0000
7	0.40	-1.738	-1.262	0.047	0.0020	0.0025	0.3575	0.2400	0.2400	0.0825
8	0.40	1.262	1.738	0.050	0.0024	0.0050	0.3250	0.2400	0.2400	0.1700
9	0.80	-0.120	0.120	0.049	0.0021	0.0025	0.3075	0.2100	0.2100	0.1225
10	1.00	-1.691	-1.309	0.049	0.0013	0.0000	0.5725	0.4525	0.4500	0.0175
11	1.40	1.303	1.697	0.048	0.0019	0.0025	0.3350	0.2325	0.2325	0.1000
12	0.40	-1.977	-1.023	0.098	0.0020	0.0000	0.8000	0.7200	0.7200	0.4775
13	0.40	1.023	1.977	0.098	0.0023	0.0000	0.6925	0.5775	0.5775	0.4875
14	0.80	-0.241	0.241	0.101	0.0021	0.0000	0.8025	0.6875	0.6875	0.5025
15	1.00	-1.882	-1.118	0.100	0.0014	0.0000	0.9775	0.9500	0.9450	0.4975
16	1.40	1.109	1.891	0.097	0.0020	0.0000	0.8050	0.7025	0.7025	0.4600

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.40	-1.500	-1.500	0.001	0.0018	0.0275	0.0525	0.0425	0.0425	0.0075
2	0.40	1.500	1.500	-0.003	0.0023	0.0550	0.0425	0.0500	0.0500	0.0325
3	0.80	0.000	0.000	0.000	0.0021	0.0575	0.0450	0.0525	0.0525	0.0250
4	1.00	-1.500	-1.500	-0.001	0.0014	0.0425	0.0475	0.0400	0.0375	0.0000
5	1.40	1.500	1.500	0.004	0.0019	0.0425	0.0600	0.0525	0.0525	0.0050
6	1.40	-1.500	-1.500	0.002	0.0012	0.0600	0.0525	0.0675	0.0375	0.0000
7	0.40	-1.738	-1.262	0.048	0.0020	0.0050	0.3750	0.2450	0.2450	0.0875
8	0.40	1.262	1.738	0.050	0.0024	0.0025	0.3250	0.2375	0.2375	0.1725
9	0.80	-0.120	0.120	0.050	0.0021	0.0025	0.3225	0.2150	0.2150	0.1250
10	1.00	-1.691	-1.309	0.050	0.0013	0.0000	0.5900	0.4450	0.4450	0.0225
11	1.40	1.303	1.697	0.048	0.0020	0.0025	0.3325	0.2250	0.2250	0.1050
12	0.40	-1.977	-1.023	0.099	0.0020	0.0000	0.8025	0.7250	0.7250	0.4900
13	0.40	1.023	1.977	0.098	0.0023	0.0000	0.6900	0.5950	0.5950	0.4975
14	0.80	-0.241	0.241	0.102	0.0021	0.0000	0.8100	0.6875	0.6875	0.5125
15	1.00	-1.882	-1.118	0.100	0.0014	0.0000	0.9800	0.9550	0.9475	0.4975
16	1.40	1.109	1.891	0.098	0.0020	0.0000	0.8050	0.7175	0.7175	0.4550

Table 2
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=250, n-F=250, d-T=0.5

CATSIB with Regression Correction

item #	a	b-R	b-F	β	SE(β)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.40	-1.500	-1.500	0.000	0.0019	0.0375	0.0600	0.0325	0.0325	0.0075
2	0.40	1.500	1.500	-0.002	0.0022	0.0475	0.0350	0.0475	0.0475	0.0250
3	0.80	0.000	0.000	0.000	0.0021	0.0600	0.0425	0.0375	0.0375	0.0075
4	1.00	-1.500	-1.500	0.000	0.0015	0.0450	0.0700	0.0650	0.0650	0.0025
5	1.40	1.500	1.500	0.002	0.0019	0.0425	0.0325	0.0325	0.0325	0.0025
6	1.40	-1.500	-1.500	0.001	0.0013	0.0475	0.0575	0.0600	0.0500	0.0000
7	0.40	-1.739	-1.261	0.048	0.0021	0.0075	0.3500	0.2300	0.2300	0.1050
8	0.40	1.261	1.739	0.050	0.0024	0.0050	0.2875	0.2100	0.2100	0.1650
9	0.80	-0.122	0.122	0.050	0.0023	0.0075	0.3225	0.2375	0.2375	0.1200
10	1.00	-1.691	-1.309	0.050	0.0014	0.0000	0.5500	0.3975	0.3975	0.0325
11	1.40	1.305	1.695	0.046	0.0022	0.0025	0.3225	0.2225	0.2225	0.1225
12	0.40	-1.978	-1.022	0.100	0.0019	0.0000	0.8300	0.7125	0.7125	0.4750
13	0.40	1.022	1.978	0.099	0.0023	0.0000	0.6850	0.5625	0.5625	0.4825
14	0.80	-0.244	0.244	0.104	0.0022	0.0000	0.7850	0.6775	0.6775	0.5450
15	1.00	-1.881	-1.119	0.102	0.0014	0.0000	0.9725	0.9525	0.9525	0.5300
16	1.40	1.113	1.887	0.097	0.0020	0.0000	0.7750	0.6675	0.6675	0.4725

CATSIB without Regression Correction

item #	a	b-R	b-F	β	SE(β)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.40	-1.500	-1.500	0.003	0.0018	0.0375	0.0700	0.0375	0.0375	0.0075
2	0.40	1.500	1.500	0.000	0.0022	0.0425	0.0400	0.0425	0.0425	0.0275
3	0.80	0.000	0.000	0.005	0.0021	0.0450	0.0650	0.0375	0.0375	0.0150
4	1.00	-1.500	-1.500	0.003	0.0015	0.0425	0.0825	0.0600	0.0600	0.0025
5	1.40	1.500	1.500	0.005	0.0019	0.0350	0.0450	0.0325	0.0325	0.0050
6	1.40	-1.500	-1.500	0.004	0.0013	0.0350	0.0750	0.0650	0.0550	0.0000
7	0.40	-1.739	-1.261	0.050	0.0020	0.0025	0.3600	0.2550	0.2550	0.1100
8	0.40	1.261	1.739	0.052	0.0024	0.0025	0.3150	0.2150	0.2150	0.1650
9	0.80	-0.122	0.122	0.056	0.0023	0.0075	0.3750	0.2675	0.2675	0.1650
10	1.00	-1.691	-1.309	0.053	0.0014	0.0000	0.6025	0.4450	0.4400	0.0425
11	1.40	1.305	1.695	0.050	0.0022	0.0025	0.3450	0.2625	0.2625	0.1375
12	0.40	-1.978	-1.022	0.102	0.0019	0.0000	0.8400	0.7525	0.7525	0.5300
13	0.40	1.022	1.978	0.102	0.0023	0.0000	0.7000	0.5875	0.5875	0.5175
14	0.80	-0.244	0.244	0.109	0.0022	0.0000	0.8200	0.7350	0.7350	0.6075
15	1.00	-1.881	-1.119	0.105	0.0014	0.0000	0.9850	0.9625	0.9625	0.5600
16	1.40	1.113	1.887	0.101	0.0020	0.0000	0.8150	0.7075	0.7075	0.5225

Table 3
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=250, n-F=250, d-T=1.0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\bar{\beta}}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.001	0.0021	0.0300	0.0550	0.0275	0.0275	0.015
2	0.4	1.500	1.500	0.000	0.0025	0.0375	0.0500	0.0425	0.0425	0.0425
3	0.8	0.000	0.000	-0.004	0.0024	0.0450	0.0325	0.0500	0.0500	0.0200
4	1.0	-1.500	-1.500	0.003	0.0017	0.0325	0.0775	0.0650	0.0650	0.0025
5	1.4	1.500	1.500	0.000	0.0021	0.0425	0.0325	0.0350	0.0350	0.0100
6	1.4	-1.500	-1.500	0.001	0.0015	0.0525	0.0725	0.0775	0.0500	0.0000
7	0.4	-1.741	-1.259	0.050	0.0023	0.0025	0.3275	0.2275	0.2275	0.1275
8	0.4	1.259	1.741	0.052	0.0027	0.0050	0.2550	0.1725	0.1725	0.1675
9	0.8	-0.127	0.127	0.051	0.0026	0.0125	0.2925	0.2075	0.2075	0.1775
10	1.0	-1.689	-1.311	0.052	0.0015	0.0000	0.4950	0.3850	0.3850	0.0675
11	1.4	1.310	1.690	0.043	0.0024	0.0075	0.2700	0.1825	0.1825	0.1175
12	0.4	-1.983	-1.017	0.101	0.0022	0.0000	0.7600	0.6100	0.6100	0.4925
13	0.4	1.017	1.983	0.101	0.0026	0.0000	0.6025	0.4725	0.4725	0.4650
14	0.8	-0.254	0.254	0.109	0.0025	0.0000	0.7275	0.6250	0.6250	0.5725
15	1.0	-1.878	-1.122	0.101	0.0016	0.0000	0.9325	0.8875	0.8875	0.5100
16	1.4	1.122	1.878	0.094	0.0023	0.0000	0.6900	0.5925	0.5925	0.4675

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\bar{\beta}}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.007	0.0020	0.0275	0.0775	0.0375	0.0375	0.0075
2	0.4	1.500	1.500	0.005	0.0025	0.0350	0.0525	0.0375	0.0375	0.0350
3	0.8	0.000	0.000	0.008	0.0023	0.0300	0.0725	0.0375	0.0375	0.0250
4	1.0	-1.500	-1.500	0.009	0.0016	0.0250	0.1025	0.0700	0.0675	0.0025
5	1.4	1.500	1.500	0.008	0.0021	0.0200	0.0650	0.0425	0.0425	0.0200
6	1.4	-1.500	-1.500	0.008	0.0014	0.0300	0.1050	0.0800	0.0675	0.0025
7	0.4	-1.741	-1.259	0.055	0.0022	0.0025	0.3650	0.2425	0.2425	0.1475
8	0.4	1.259	1.741	0.058	0.0026	0.0025	0.3075	0.2025	0.2025	0.1900
9	0.8	-0.127	0.127	0.063	0.0025	0.0050	0.3725	0.2625	0.2625	0.2300
10	1.0	-1.689	-1.311	0.058	0.0015	0.0000	0.5550	0.4425	0.4425	0.1000
11	1.4	1.310	1.690	0.052	0.0024	0.0050	0.3400	0.2425	0.2425	0.1675
12	0.4	-1.983	-1.017	0.106	0.0022	0.0000	0.8200	0.6825	0.6825	0.5325
13	0.4	1.017	1.983	0.106	0.0026	0.0000	0.6400	0.5350	0.5350	0.5175
14	0.8	-0.254	0.254	0.120	0.0024	0.0000	0.8075	0.7250	0.7250	0.6775
15	1.0	-1.878	-1.122	0.108	0.0016	0.0000	0.9575	0.9325	0.9325	0.5825
16	1.4	1.122	1.878	0.102	0.0022	0.0000	0.7500	0.6525	0.6525	0.5325

Table 4
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=250, d-T=0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	-0.002	0.0016	0.0675	0.0400	0.0450	0.0450	0.0025
2	0.4	1.500	1.500	-0.002	0.0020	0.0550	0.0500	0.0600	0.0600	0.0150
3	0.8	0.000	0.000	0.001	0.0019	0.0425	0.0500	0.0475	0.0475	0.0100
4	1.0	-1.500	-1.500	0.000	0.0012	0.0525	0.0550	0.0500	0.0350	0.0000
5	1.4	1.500	1.500	-0.003	0.0015	0.0450	0.0300	0.0250	0.0250	0.0000
6	1.4	-1.500	-1.500	0.000	0.0010	0.0625	0.0300	0.0375	0.0100	0.0000
7	0.4	-1.656	-1.188	0.052	0.0018	0.0025	0.4450	0.3200	0.3200	0.0875
8	0.4	1.338	1.824	0.048	0.0021	0.0025	0.3525	0.2625	0.2625	0.1125
9	0.8	-0.080	0.160	0.054	0.0019	0.0000	0.4125	0.3100	0.3100	0.1025
10	1.0	-1.622	-1.256	0.047	0.0013	0.0000	0.5925	0.4725	0.4425	0.0175
11	1.4	1.359	1.782	0.049	0.0016	0.0000	0.4375	0.3150	0.3150	0.0650
12	0.4	-1.806	-0.888	0.101	0.0017	0.0000	0.9125	0.8375	0.8375	0.5275
13	0.4	1.168	2.164	0.099	0.0019	0.0000	0.8325	0.7125	0.7125	0.5100
14	0.8	-0.161	0.322	0.105	0.0018	0.0000	0.8950	0.8100	0.8100	0.5200
15	1.0	-1.734	-1.032	0.102	0.0015	0.0000	0.9800	0.9725	0.9700	0.5125
16	1.4	1.198	2.104	0.100	0.0017	0.0000	0.9200	0.8650	0.8650	0.5050

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	-0.002	0.0016	0.0675	0.0350	0.0475	0.0475	0.0025
2	0.4	1.500	1.500	-0.002	0.0020	0.0550	0.0575	0.0625	0.0625	0.0125
3	0.8	0.000	0.000	0.001	0.0019	0.0400	0.0550	0.0375	0.0375	0.0100
4	1.0	-1.500	-1.500	0.000	0.0012	0.0625	0.0575	0.0525	0.0325	0.0000
5	1.4	1.500	1.500	-0.003	0.0015	0.0375	0.0300	0.0275	0.0275	0.0000
6	1.4	-1.500	-1.500	0.000	0.0010	0.0650	0.0350	0.0375	0.0075	0.0000
7	0.4	-1.656	-1.188	0.053	0.0018	0.0000	0.4550	0.3350	0.3350	0.0900
8	0.4	1.338	1.824	0.049	0.0021	0.0000	0.3425	0.2600	0.2600	0.1200
9	0.8	-0.080	0.160	0.054	0.0019	0.0000	0.4150	0.3200	0.3200	0.1050
10	1.0	-1.622	-1.256	0.047	0.0013	0.0000	0.5875	0.4575	0.4275	0.0200
11	1.4	1.359	1.782	0.049	0.0016	0.0000	0.4225	0.3150	0.3150	0.0700
12	0.4	-1.806	-0.888	0.101	0.0017	0.0000	0.9125	0.8325	0.8325	0.5125
13	0.4	1.168	2.164	0.098	0.0019	0.0000	0.8275	0.7225	0.7225	0.4925
14	0.8	-0.161	0.322	0.104	0.0018	0.0000	0.9000	0.8125	0.8125	0.5225
15	1.0	-1.734	-1.032	0.103	0.0015	0.0000	0.9800	0.9675	0.9625	0.5375
16	1.4	1.198	2.104	0.100	0.0017	0.0000	0.9275	0.8550	0.8550	0.5100

Table 5
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=250, d-T=0.5

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\bar{\beta}}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	-0.002	0.0017	0.0600	0.0400	0.0475	0.0475	0.0025
2	0.4	1.500	1.500	-0.001	0.0020	0.0475	0.0500	0.0425	0.0425	0.0075
3	0.8	0.000	0.000	0.001	0.0018	0.0400	0.0500	0.0400	0.0400	0.0050
4	1.0	-1.500	-1.500	0.002	0.0012	0.0500	0.0675	0.0575	0.0400	0.0000
5	1.4	1.500	1.500	-0.004	0.0017	0.0375	0.0475	0.0400	0.0400	0.0025
6	1.4	-1.500	-1.500	0.002	0.0009	0.0475	0.0400	0.0425	0.0100	0.0000
7	0.4	-1.656	-1.188	0.053	0.0018	0.0000	0.4500	0.3200	0.3200	0.0875
8	0.4	1.338	1.824	0.048	0.0022	0.0100	0.3525	0.2400	0.2400	0.1250
9	0.8	-0.081	0.162	0.053	0.0019	0.0025	0.3825	0.2925	0.2925	0.0925
10	1.0	-1.622	-1.256	0.050	0.0013	0.0025	0.6550	0.5125	0.4800	0.0350
11	1.4	1.361	1.778	0.046	0.0017	0.0000	0.3400	0.2725	0.2725	0.0625
12	0.4	-1.807	-0.886	0.104	0.0018	0.0000	0.8950	0.8250	0.8250	0.5675
13	0.4	1.167	2.166	0.097	0.0020	0.0000	0.7825	0.6925	0.6925	0.4800
14	0.8	-0.163	0.326	0.107	0.0019	0.0000	0.8900	0.8025	0.8025	0.5750
15	1.0	-1.734	-1.032	0.105	0.0014	0.0000	0.9900	0.9850	0.9800	0.5675
16	1.4	1.204	2.092	0.092	0.0017	0.0000	0.8525	0.7475	0.7475	0.4125

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\bar{\beta}}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.000	0.0017	0.0550	0.0500	0.0475	0.0475	0.0050
2	0.4	1.500	1.500	0.002	0.0020	0.0375	0.0550	0.0575	0.0575	0.0075
3	0.8	0.000	0.000	0.007	0.0018	0.0275	0.0675	0.0325	0.0325	0.0050
4	1.0	-1.500	-1.500	0.005	0.0012	0.0325	0.0850	0.0675	0.0550	0.0000
5	1.4	1.500	1.500	0.000	0.0017	0.0425	0.0475	0.0425	0.0425	0.0050
6	1.4	-1.500	-1.500	0.005	0.0010	0.0325	0.0600	0.0450	0.0150	0.0000
7	0.4	-1.656	-1.188	0.055	0.0018	0.0000	0.4800	0.3600	0.3600	0.1200
8	0.4	1.338	1.824	0.051	0.0022	0.0050	0.3900	0.2675	0.2675	0.1375
9	0.8	-0.081	0.162	0.059	0.0019	0.0000	0.4450	0.3500	0.3500	0.1250
10	1.0	-1.622	-1.256	0.053	0.0013	0.0000	0.6900	0.5675	0.5250	0.0400
11	1.4	1.361	1.778	0.049	0.0017	0.0000	0.4050	0.2800	0.2800	0.0700
12	0.4	-1.807	-0.886	0.108	0.0018	0.0000	0.9125	0.8500	0.8500	0.5800
13	0.4	1.167	2.166	0.100	0.0020	0.0000	0.8175	0.7350	0.7350	0.5025
14	0.8	-0.163	0.326	0.112	0.0019	0.0000	0.9150	0.8675	0.8675	0.6275
15	1.0	-1.734	-1.032	0.109	0.0014	0.0000	0.9900	0.9875	0.9850	0.6000
16	1.4	1.204	2.092	0.095	0.0017	0.0000	0.8725	0.7675	0.7675	0.4500

Table 6
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=250, d-T=1.0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	-0.001	0.0019	0.0650	0.0375	0.0450	0.0450	0.0025
2	0.4	1.500	1.500	-0.001	0.0022	0.0500	0.0350	0.0350	0.0350	0.0125
3	0.8	0.000	0.000	0.000	0.0022	0.0750	0.0500	0.0750	0.0750	0.0225
4	1.0	-1.500	-1.500	0.004	0.0014	0.0500	0.0875	0.0800	0.0650	0.0000
5	1.4	1.500	1.500	-0.005	0.0019	0.0500	0.0575	0.0500	0.0500	0.0050
6	1.4	-1.500	-1.500	0.004	0.0011	0.0400	0.0725	0.0550	0.0200	0.0000
7	0.4	-1.657	-1.186	0.053	0.0021	0.0025	0.4250	0.2850	0.2850	0.1550
8	0.4	1.337	1.826	0.047	0.0024	0.0000	0.2925	0.2025	0.2025	0.1450
9	0.8	-0.084	0.168	0.053	0.0021	0.0025	0.3175	0.2300	0.2300	0.1450
10	1.0	-1.622	-1.256	0.055	0.0015	0.0025	0.6600	0.5400	0.5250	0.0600
11	1.4	1.367	1.766	0.040	0.0020	0.0050	0.3050	0.2125	0.2125	0.0600
12	0.4	-1.810	-0.880	0.107	0.0020	0.0000	0.8650	0.7475	0.7475	0.5400
13	0.4	1.166	2.168	0.100	0.0022	0.0000	0.7275	0.6275	0.6275	0.5250
14	0.8	-0.168	0.336	0.108	0.0022	0.0000	0.7975	0.6950	0.6950	0.5700
15	1.0	-1.736	-1.028	0.112	0.0016	0.0000	0.9925	0.9800	0.9800	0.6275
16	1.4	1.217	2.066	0.080	0.0020	0.0000	0.6275	0.5425	0.5425	0.3275

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.006	0.0019	0.0425	0.0475	0.0500	0.0500	0.0050
2	0.4	1.500	1.500	0.006	0.0022	0.0225	0.0500	0.0325	0.0325	0.0175
3	0.8	0.000	0.000	0.012	0.0021	0.0375	0.1000	0.0750	0.0750	0.0275
4	1.0	-1.500	-1.500	0.010	0.0014	0.0275	0.1500	0.0900	0.0850	0.0050
5	1.4	1.500	1.500	0.003	0.0019	0.0375	0.0775	0.0525	0.0525	0.0025
6	1.4	-1.500	-1.500	0.010	0.0011	0.0200	0.1325	0.0750	0.0275	0.0000
7	0.4	-1.657	-1.186	0.060	0.0021	0.0025	0.4900	0.3500	0.3500	0.1600
8	0.4	1.337	1.826	0.053	0.0023	0.0025	0.3475	0.2400	0.2400	0.1625
9	0.8	-0.084	0.168	0.065	0.0021	0.0000	0.4375	0.3200	0.3200	0.2075
10	1.0	-1.622	-1.256	0.063	0.0015	0.0000	0.7350	0.6525	0.6450	0.1050
11	1.4	1.367	1.766	0.047	0.0019	0.0025	0.3450	0.2625	0.2625	0.0700
12	0.4	-1.810	-0.880	0.113	0.0020	0.0000	0.8925	0.8175	0.8175	0.5900
13	0.4	1.166	2.168	0.106	0.0022	0.0000	0.7675	0.6675	0.6675	0.5750
14	0.8	-0.168	0.336	0.119	0.0021	0.0000	0.8950	0.7950	0.7950	0.6650
15	1.0	-1.736	-1.028	0.119	0.0016	0.0000	1.0000	0.9900	0.9900	0.7075
16	1.4	1.217	2.066	0.088	0.0019	0.0000	0.7100	0.6225	0.6225	0.3850

Table 7
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=500, d-T=0.0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.000	0.0014	0.0625	0.0725	0.0725	0.0725	0.0000
2	0.4	1.500	1.500	0.001	0.0015	0.0525	0.0575	0.0500	0.0500	0.0000
3	0.8	0.000	0.000	-0.001	0.0015	0.0575	0.0425	0.0650	0.0650	0.0025
4	1.0	-1.500	-1.500	-0.001	0.0010	0.0525	0.0450	0.0625	0.0075	0.0000
5	1.4	1.500	1.500	-0.001	0.0014	0.0500	0.0475	0.0450	0.0450	0.0000
6	1.4	-1.500	-1.500	0.000	0.0008	0.0425	0.0400	0.0450	0.0000	0.0000
7	0.4	-1.738	-1.262	0.050	0.0013	0.0000	0.5500	0.4450	0.4450	0.0250
8	0.4	1.262	1.738	0.050	0.0016	0.0000	0.4700	0.3425	0.3425	0.0675
9	0.8	-0.120	0.120	0.052	0.0015	0.0000	0.5400	0.4025	0.4025	0.0500
10	1.0	-1.691	-1.309	0.050	0.0010	0.0000	0.8225	0.7375	0.5000	0.0050
11	1.4	1.303	1.697	0.050	0.0013	0.0050	0.5975	0.4600	0.4600	0.0225
12	0.4	-1.977	-1.023	0.103	0.0013	0.0000	0.9800	0.9625	0.9625	0.5450
13	0.4	1.023	1.977	0.101	0.0017	0.0000	0.9125	0.8750	0.8750	0.4925
14	0.8	-0.241	0.241	0.102	0.0015	0.0000	0.9625	0.9250	0.9250	0.5250
15	1.0	-1.882	-1.118	0.100	0.0010	0.0000	1.0000	1.0000	0.9950	0.5275
16	1.4	1.109	1.891	0.098	0.0014	0.0000	0.9650	0.9350	0.9350	0.4525

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.000	0.0014	0.0650	0.0700	0.0725	0.0725	0.0000
2	0.4	1.500	1.500	0.001	0.0016	0.0500	0.0625	0.0525	0.0525	0.0000
3	0.8	0.000	0.000	0.000	0.0015	0.0600	0.0425	0.0675	0.0675	0.0025
4	1.0	-1.500	-1.500	-0.001	0.0010	0.0625	0.0525	0.0600	0.0125	0.0000
5	1.4	1.500	1.500	-0.001	0.0014	0.0525	0.0450	0.0425	0.0425	0.0000
6	1.4	-1.500	-1.500	0.000	0.0008	0.0475	0.0525	0.0525	0.0000	0.0000
7	0.4	-1.738	-1.262	0.050	0.0013	0.0000	0.5625	0.4375	0.4375	0.0225
8	0.4	1.262	1.738	0.050	0.0016	0.0000	0.4700	0.3575	0.3575	0.0650
9	0.8	-0.120	0.120	0.052	0.0015	0.0000	0.5250	0.4125	0.4125	0.0575
10	1.0	-1.691	-1.309	0.050	0.0010	0.0000	0.8050	0.7425	0.4950	0.0050
11	1.4	1.303	1.697	0.050	0.0013	0.0050	0.6075	0.4675	0.4675	0.0250
12	0.4	-1.977	-1.023	0.103	0.0014	0.0000	0.9800	0.9550	0.9550	0.5550
13	0.4	1.023	1.977	0.101	0.0017	0.0000	0.9150	0.8650	0.8650	0.4825
14	0.8	-0.241	0.241	0.102	0.0015	0.0000	0.9600	0.9325	0.9325	0.5350
15	1.0	-1.882	-1.118	0.100	0.0010	0.0000	1.0000	1.0000	0.9950	0.5125
16	1.4	1.109	1.891	0.097	0.0014	0.0000	0.9750	0.9425	0.9425	0.4550

Table 8
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=500, d-T=0.5

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.000	0.0015	0.0625	0.0550	0.0625	0.0625	0.0000
2	0.4	1.500	1.500	0.003	0.0016	0.0350	0.0525	0.0450	0.0450	0.0025
3	0.8	0.000	0.000	-0.001	0.0016	0.0575	0.0425	0.0450	0.0450	0.0050
4	1.0	-1.500	-1.500	-0.001	0.0010	0.0600	0.0500	0.0525	0.0100	0.0000
5	1.4	1.500	1.500	-0.001	0.0014	0.0475	0.0550	0.0425	0.0425	0.0000
6	1.4	-1.500	-1.500	0.002	0.0009	0.0400	0.0750	0.0600	0.0000	0.0000
7	0.4	-1.739	-1.261	0.050	0.0014	0.0000	0.5600	0.4300	0.4300	0.0350
8	0.4	1.261	1.739	0.048	0.0017	0.0000	0.4575	0.3250	0.3250	0.0725
9	0.8	-0.122	0.122	0.052	0.0015	0.0000	0.5250	0.4000	0.4000	0.0500
10	1.0	-1.691	-1.309	0.050	0.0010	0.0000	0.8225	0.7300	0.5175	0.0075
11	1.4	1.305	1.695	0.050	0.0014	0.0050	0.5325	0.4225	0.4225	0.0350
12	0.4	-1.978	-1.022	0.102	0.0014	0.0000	0.9700	0.9450	0.9450	0.5325
13	0.4	1.022	1.978	0.102	0.0017	0.0000	0.9200	0.8750	0.8750	0.5350
14	0.8	-0.244	0.244	0.103	0.0015	0.0000	0.9750	0.9250	0.9250	0.5325
15	1.0	-1.881	-1.119	0.100	0.0010	0.0000	1.0000	1.0000	0.9975	0.4750
16	1.4	1.113	1.887	0.097	0.0015	0.0000	0.9700	0.9275	0.9275	0.4625

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.003	0.0015	0.0450	0.0700	0.0675	0.0675	0.0000
2	0.4	1.500	1.500	0.006	0.0015	0.0200	0.0575	0.0475	0.0475	0.0025
3	0.8	0.000	0.000	0.005	0.0015	0.0250	0.0700	0.0550	0.0550	0.0025
4	1.0	-1.500	-1.500	0.002	0.0010	0.0450	0.0725	0.0550	0.0125	0.0000
5	1.4	1.500	1.500	0.003	0.0014	0.0350	0.0700	0.0400	0.0400	0.0000
6	1.4	-1.500	-1.500	0.004	0.0009	0.0325	0.1025	0.0575	0.0000	0.0000
7	0.4	-1.739	-1.261	0.053	0.0014	0.0000	0.5850	0.4675	0.4675	0.0425
8	0.4	1.261	1.739	0.051	0.0017	0.0000	0.4825	0.3750	0.3750	0.0750
9	0.8	-0.122	0.122	0.057	0.0015	0.0000	0.5900	0.4525	0.4525	0.0625
10	1.0	-1.691	-1.309	0.053	0.0010	0.0000	0.8500	0.7925	0.5800	0.0150
11	1.4	1.305	1.695	0.053	0.0014	0.0050	0.5900	0.4850	0.4850	0.0450
12	0.4	-1.978	-1.022	0.104	0.0014	0.0000	0.9800	0.9500	0.9500	0.5650
13	0.4	1.022	1.978	0.105	0.0017	0.0000	0.9350	0.8900	0.8900	0.5700
14	0.8	-0.244	0.244	0.108	0.0015	0.0000	0.9825	0.9575	0.9575	0.5900
15	1.0	-1.881	-1.119	0.103	0.0010	0.0000	1.0000	1.0000	1.0000	0.5475
16	1.4	1.113	1.887	0.100	0.0014	0.0000	0.9725	0.9400	0.9400	0.5000

Table 9
Item Parameters, Estimated DIF, and rejection rates of CATSIB
n-R=500, n-F=500, d-T=1.0

CATSIB with Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.001	0.0016	0.0600	0.0650	0.0650	0.0650	0.0075
2	0.4	1.500	1.500	0.004	0.0018	0.0450	0.0375	0.0350	0.0350	0.0050
3	0.8	0.000	0.000	0.000	0.0016	0.0400	0.0400	0.0475	0.0475	0.0050
4	1.0	-1.500	-1.500	0.002	0.0012	0.0600	0.0700	0.0750	0.0350	0.0000
5	1.4	1.500	1.500	-0.002	0.0015	0.0525	0.0375	0.0650	0.0650	0.0025
6	1.4	-1.500	-1.500	0.003	0.0010	0.0275	0.0950	0.0625	0.0050	0.0000
7	0.4	-1.741	-1.259	0.051	0.0015	0.0000	0.4925	0.3700	0.3700	0.0525
8	0.4	1.259	1.741	0.049	0.0018	0.0025	0.3675	0.2650	0.2650	0.0875
9	0.8	-0.127	0.127	0.052	0.0016	0.0000	0.5050	0.3625	0.3625	0.0625
10	1.0	-1.689	-1.311	0.052	0.0010	0.0000	0.7625	0.6350	0.5425	0.0125
11	1.4	1.310	1.690	0.050	0.0016	0.0025	0.4700	0.3525	0.3525	0.0675
12	0.4	-1.983	-1.017	0.103	0.0015	0.0000	0.9550	0.9150	0.9150	0.5275
13	0.4	1.017	1.983	0.103	0.0018	0.0000	0.8875	0.8350	0.8350	0.5200
14	0.8	-0.254	0.254	0.107	0.0017	0.0000	0.9350	0.8825	0.8825	0.5525
15	1.0	-1.878	-1.122	0.101	0.0012	0.0000	0.9950	0.9925	0.9850	0.5175
16	1.4	1.122	1.878	0.096	0.0016	0.0000	0.9150	0.8450	0.8450	0.4400

CATSIB without Regression Correction

item #	a	b-R	b-F	$\hat{\beta}$	SE($\hat{\beta}$)	LT rr	RT rr	2T rr	MDIF rr	BDIF rr
1	0.4	-1.500	-1.500	0.007	0.0016	0.0325	0.0775	0.0700	0.0700	0.0050
2	0.4	1.500	1.500	0.009	0.0017	0.0325	0.0700	0.0425	0.0425	0.0050
3	0.8	0.000	0.000	0.011	0.0016	0.0200	0.0850	0.0600	0.0600	0.0050
4	1.0	-1.500	-1.500	0.008	0.0012	0.0275	0.1125	0.0900	0.0425	0.0025
5	1.4	1.500	1.500	0.007	0.0015	0.0300	0.0800	0.0600	0.0600	0.0000
6	1.4	-1.500	-1.500	0.009	0.0010	0.0150	0.1500	0.1050	0.0200	0.0000
7	0.4	-1.741	-1.259	0.057	0.0015	0.0000	0.5925	0.4625	0.4625	0.0750
8	0.4	1.259	1.741	0.055	0.0018	0.0025	0.4625	0.3425	0.3425	0.1050
9	0.8	-0.127	0.127	0.064	0.0016	0.0000	0.6275	0.5150	0.5150	0.1025
10	1.0	-1.689	-1.311	0.059	0.0010	0.0000	0.8600	0.7650	0.6775	0.0275
11	1.4	1.310	1.690	0.058	0.0016	0.0025	0.5950	0.4275	0.4275	0.0875
12	0.4	-1.983	-1.017	0.109	0.0015	0.0000	0.9825	0.9500	0.9500	0.5925
13	0.4	1.017	1.983	0.108	0.0018	0.0000	0.9250	0.8850	0.8850	0.5800
14	0.8	-0.254	0.254	0.117	0.0017	0.0000	0.9775	0.9425	0.9425	0.6925
15	1.0	-1.878	-1.122	0.107	0.0012	0.0000	1.0000	1.0000	0.9975	0.6175
16	1.4	1.122	1.878	0.104	0.0016	0.0000	0.9550	0.9125	0.9125	0.5525

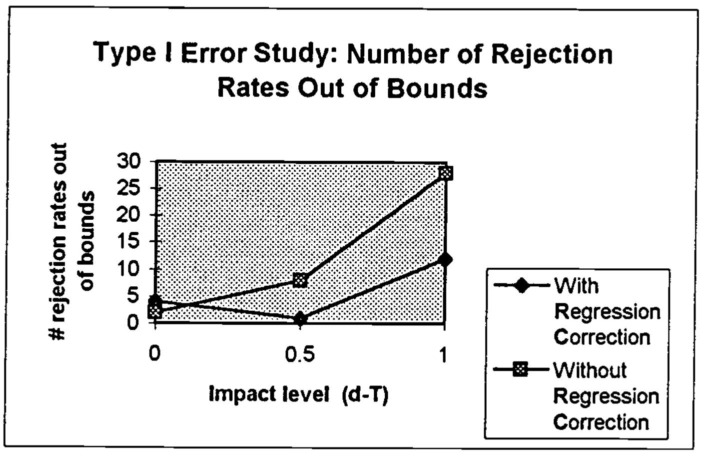
Table 10
Summary of Type I Error Rates

	With Regression Correction			Without Regression Correction		
n-R, n-F	**d-T=0**	**d-T=0.5**	**d-T=1.0**	**d-T=0**	**d-T=0.5**	**d-T=1.0**
LT rr						
250, 250	0.0442	0.0467	0.0400	0.0475	0.0396	0.0279
500, 250	0.0542	0.0471	0.0550	0.0546	0.0379	0.0313
500, 500	0.0529	0.0504	0.0475	0.0563	0.0338	0.0263
mean	**0.0504**	**0.0481**	**0.0475**	**0.0528**	**0.0371**	**0.0285**
RT rr						
250, 250	0.0517	0.0496	0.0533	0.0500	0.0629	0.0792
500, 250	0.0425	0.0492	0.0567	0.0450	0.0608	0.0929
500, 500	0.0508	0.0550	0.0575	0.0542	0.0738	0.0958
mean	**0.0483**	**0.0513**	**0.0558**	**0.0497**	**0.0658**	**0.0893**
2T rr						
250, 250	0.0508	0.0458	0.0496	0.0508	0.0458	0.0508
500, 250	0.0442	0.0450	0.0567	0.0442	0.0488	0.0625
500, 500	0.0567	0.0513	0.0583	0.0579	0.0538	0.0713
mean	**0.0506**	**0.0474**	**0.0549**	**0.051**	**0.0494**	**0.0615**

Table 11
Type I Error Study: Number of Rejection Rates Out of Bounds

n-R, n-F	d-T=0	d-T=0.5	d-T=1.0
With Regression Correction			
250, 250	1	0	4
500, 250	1	0	5
500, 500	2	1	3
Total	4	1	12
Without Regression Correction			
250, 250	1	2	8
500, 250	0	2	10
500, 500	1	4	10
Total	2	8	28

Figure 1



27

Table 12
Summary of Power Rates

	Sample Sizes for n-R and n-F		
	250,250	500,250	500,500
RT rr for d-T=0			
MDIF	38%	45%	59%
HDIF	82%	91%	97%
RT rr for d-T=0.5			
MDIF	37%	44%	58%
HDIF	81%	88%	97%
Rt rr for d-T=1.0			
MDIF	33%	40%	52%
HDIF	74%	80%	94%
2T rr for d-T=0			
MDIF	28%	34%	47%
HDIF	73%	84%	94%
2T rr for d-T=0.5			
MDIF	26%	33%	46%
HDIF	71%	81%	94%
2T rr for d-T=1.0			
MDIF	24%	29%	40%
HDIF	64%	72%	89%

AERA CONFERENCE 1997

TM026775

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Validation of CATSIB to Investigate DIF of CAT Data

Author(s): Ratna Nandakumar and Louis Roussos

Corporate Source: University of Delaware & Law School Admission Council.

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

[✓] Check here
For Level 1 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all Level 2 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

[ ] Check here
For Level 2 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Ratna Nandakumar

Printed Name/Position/Title: Ratna Nandakumar, Associate Professor, Univ of Delaware

Organization/Address: Univ of Delaware, 213 Willard Hall, Newark, DE 19716

Telephone: 302-831-1635

FAX: 302-831-4445

E-Mail Address: nandakum@udel.edu

Date: 3/27/97

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Assessment and Evaluation
210 O'Boyle Hall
The Catholic University of America
Washington, DC 20064

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2d Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com